**Komentář k oddílu XI.1 – Lokální extrémy – podmínky druhého řádu**

**K definicím souvisejících pojmů**

- Lokální extrémy, tj. lokální maximum a lokální minimum, byly definovány pro funkce jedné proměnné v oddílu IV.7 a pro funkce více proměnných v oddílu V.2.

  Význam je přirozený – funkce $f$ má v bodě $\boldsymbol{x}$ lokální maximum, pokud je hodnota v bodě $\boldsymbol{x}$ největší na nějakém okolí bodu $\boldsymbol{x}$.

  Podobně, $f$ má v bodě $\boldsymbol{x}$ lokální minimum, pokud je hodnota v bodě $\boldsymbol{x}$ nejmenší na nějakém okolí bodu $\boldsymbol{x}$.

- Novými pojmy jsou ostré lokální maximum a ostré lokální minimum.

  $f$ má v bodě $\boldsymbol{x}$ ostré lokální maximum, pokud existuje okolí bodu $\boldsymbol{x}$, na kterém je v bodě $\boldsymbol{x}$ největší hodnota a navíc v ostatních bodech je hodnota menší.

  Analogicky, $f$ má v bodě $\boldsymbol{x}$ ostré lokální minimum, pokud existuje okolí bodu $\boldsymbol{x}$, na kterém je v bodě $\boldsymbol{x}$ nejmenší hodnota a navíc v ostatních bodech je hodnota větší.

  Rozdíl oproti lokálnímu maximu a minimu ilustrují následující příklady:

  Uvažme dvě funkce dvou proměnných

  $$f(x,y) = x^2 + y^2 \quad \text{a} \quad g(x,y) = (x+y)^2, \quad [x,y] \in \mathbf{R}^2.$$

  Funkce $f$ má v bodě $[0,0]$ ostré lokální minimum – $f(0,0) = 0$ a $f(x,y) > 0$ pro $[x,y] \in \mathbf{R}^2 \setminus \{[0,0]\}$.

  Naproti tomu funkce $g$ má v bodě $[0,0]$ lokální minimum, které není ostré.

  To, že jde o lokální minimum plyne z toho, že $g(0,0) = 0$ a $g(x,y) \geq 0$ pro každé $[x,y] \in \mathbf{R}^2$.

  To, že není ostré, plyne z toho, že $g$ nabývá hodnoty 0 v každém bodě přímky $y = -x$. Nemůžeme tedy najít $\delta > 0$, aby $g(x,y) > 0$ pro každé $[x,y] \in B([0,0],\delta) \setminus \{[0,0]\}$ (bod $[\frac{\delta}{2}, -\frac{\delta}{2}]$ patří do $B([0,0],\delta) \setminus \{[0,0]\}$, ale $g(\frac{\delta}{2}, -\frac{\delta}{2}) = 0$).

- Pokud $G \subset \mathbf{R}^n$ je otevřená, $f \in C^1(G)$ a funkce $f$ má v bodě $\boldsymbol{a} \in G$ lokální extrém, pak $\nabla f(\boldsymbol{a}) = 0$ – to plyne z Věty V.6.

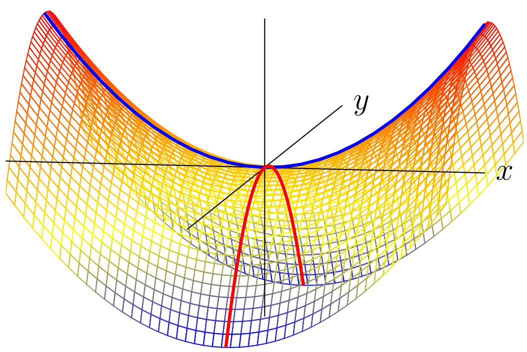
  Proto se zavádí pojem stacionárního bodu – to je takový bod $\boldsymbol{a} \in G$, pro který platí $\nabla f(\boldsymbol{a}) = 0$.

  Tedy z Věty V.6 plyne, že bod lokálního extrému pro funkci třídy $C^1$ je nutně bod stacionární.

  Obrácená implikace neplatí, nulovost parciálních derivací nezaručí, že je v daném bodě lokální extrém; jinými slovy ve stacionárním bodě nemusí být lokální extrém.

  Tato situace si zaslouží zvláštní název – stacionárnímu bodu, který není bodem lokálního extrému, se říká sedlový bod.

  Příkladem sedlového bodu je například bod 0 pro funkci $f(x) = x^3$, nebo bod $[0, 0]$ pro funkci dvou proměnných $g(x, y) = x^2 - y^2$.

  Na obrázku je znázorněn sedlový bod $[0, 0]$ pro funkci $f(x, y) = x^2 - y^2$. Kromě grafu funkce jsou vyznačeny části grafu pro $y = 0$ (modře, $g(x, 0) = x^2$) a pro $x = 0$ (červeně, $g(0, y) = -y^2$). Ty názorně vytvářejí sedlo.

**Věta XI.1 a její důkaz:**

- Pokud funkce $f$ třídy $C^2$ má v bodě $\boldsymbol{a}$ lokální extrém, víme podle Věty V.6, že $\nabla f(\boldsymbol{a}) = \boldsymbol{o}$. Věta XI.1 dává podrobnější informaci s použitím Hessovy matice. Pokud je bodě $\boldsymbol{a}$ lokální maximum, je Hessova matice negativně semidefinitní; pokud je v bodě $\boldsymbol{a}$ lokální maximum, je Hessova matice pozitivně semidefinitní. Tyto podmínky jsou nutné, nikoli postačující, jak podrobněji vysvětlíme níže.

- Dokážeme bod (i), bod (ii) je zcela analogický (případně lze dokázat aplikací bodu (i) na funkci $-f$).

- Dokažme bod (i) nejprve pro $n = 1$, tj. pro funkci jedné proměnné.

  Nechť $f \in C^2(a-r, a+r)$ a $f$ má v bodě $a$ lokální maximum. Pak $f''(a) \leq 0$.

  Předpokládejme pro spor, že $f''(a) > 0$. Protože $f$ má v bodě $a$ lokální maximum, platí $f'(a) = 0$ (Věta IV.31).

  Dále si uvědomme, že

  $$f''(a) = \lim_{x \to a} \frac{f'(x) - f'(a)}{x - a} \overset{f'(a)=0}{=} \lim_{x \to a} \frac{f'(x)}{x - a}.$$

  Protože $f''(a) > 0$, máme

  $$\lim_{x \to a} \frac{f'(x)}{x - a} > 0,$$

  tedy (podle věty o limitě a nerovnostech, viz Věta IV.5) existuje $\delta > 0$, že

  $$\forall x \in P(a, \delta) : \frac{f'(x)}{x - a} > 0.$$

  Tedy

  $$\forall x \in (a, a+\delta) : f'(x) > 0 \quad \text{and} \quad \forall x \in (a-\delta, a) : f'(x) < 0.$$

  Podle věty o znaménku derivace a monotonii nyní platí, že $f$ je klesající na $(a-\delta, a\rangle$ a rostoucí na $\langle a, a+\delta)$.

  To je ovšem spor s tím, že v bodě $a$ je lokální maximum – vyšlo nám totiž, že v bodě $a$ je ostré lokální minimum (viz následující tabulka).

  | $x$ | $a-\delta$ | | $a$ | | $a+\delta$ |
  |---|---|---|---|---|---|
  | $f'(x)$ | | $< 0$ | $0$ | $> 0$ | |
  | $f(x)$ | | $\searrow$ | ostré lok. min. | $\nearrow$ | |

- Nyní dokažme bod (i). Předpokládejme, že v bodě $\boldsymbol{a}$ je lokální maximum.

  To znamená, že existuje $\delta > 0$, že

  $$\forall \boldsymbol{x} \in B(\boldsymbol{a}, \delta) \colon f(\boldsymbol{x}) \leq f(\boldsymbol{a}).$$

Chceme ukázat, že $\nabla^2 f(\boldsymbol{a})$ je negativně semidefinitní, tj., že

$$\forall \boldsymbol{h} \in \mathbf{R}^n : \boldsymbol{h}^T \nabla^2 f(\boldsymbol{a}) \boldsymbol{h} \leq 0.$$

Zvolme tedy $\boldsymbol{h} \in \mathbf{R}^n$ libovolné. Pokud $\boldsymbol{h} = \boldsymbol{o}$, pak

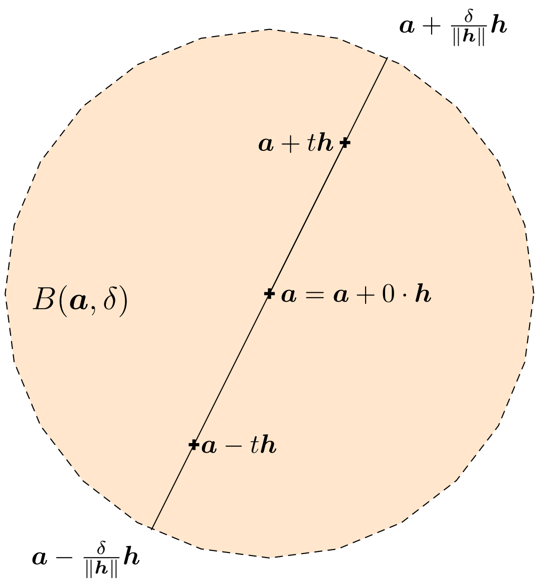
$$\boldsymbol{h}^T \nabla^2 f(\boldsymbol{a}) \boldsymbol{h} = \boldsymbol{o}^T \nabla^2 f(\boldsymbol{a}) \boldsymbol{o} = 0 \leq 0.$$

Předpokládejme tedy, že $\boldsymbol{h} \neq \boldsymbol{o}$.

Definujeme pomocnou funkci jedné proměnné předpisem

$$\varphi(t) = f(\boldsymbol{a} + t\boldsymbol{h}).$$

Pak $\varphi$ je definována na nějakém okolí 0 a má v bodě 0 lokální maximum. Ilustruje to následující obrázek:




Množina $\{\boldsymbol{a} + t\boldsymbol{h} : t \in \mathbf{R}\}$ je přímka procházející bodem $\boldsymbol{a}$ ve směru $\boldsymbol{h}$. Pro $t \in (-\frac{\delta}{\|h\|}, \frac{\delta}{\|h\|})$ platí $\boldsymbol{a} + t\boldsymbol{h} \in B(\boldsymbol{a}, \delta)$, a tedy

$$\varphi(t) = f(\boldsymbol{a} + t\boldsymbol{h}) \leq f(\boldsymbol{a}) = \varphi(0).$$

Opravdu má tedy $\varphi$ v bodě 0 lokální minimum.

Navíc je funkce $\varphi$ třídy $C^2$ na intervalu $(-\frac{\delta}{\|h\|}, \frac{\delta}{\|h\|})$ (podle věty o derivaci složené funkce – viz Věta V.14).

Tedy, podle již dokázaného případu $n = 1$ víme, že platí $\varphi''(0) \leq 0$. Zbývá spočítat $\varphi''(0)$. Použije opět Větu V.14:

$$\varphi(t) = f(a_1 + th_1, \ldots, a_n + th_n), \qquad t \in (-\tfrac{\delta}{\|h\|}, \tfrac{\delta}{\|h\|});$$

$$\varphi'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a_1 + th_1, \ldots, a_n + th_n) \cdot h_i, \qquad t \in (-\tfrac{\delta}{\|h\|}, \tfrac{\delta}{\|h\|});$$

$$\varphi''(t) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a_1 + th_1, \ldots, a_n + th_n) \cdot h_j \cdot h_i, \quad t \in (-\tfrac{\delta}{\|h\|}, \tfrac{\delta}{\|h\|});$$

$$\varphi''(0) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) \cdot h_j \cdot h_i = \boldsymbol{h}^T \nabla^2 f(\boldsymbol{a}) \boldsymbol{h}.$$

Je tedy $\boldsymbol{h}^T \nabla^2 f(\boldsymbol{a}) \boldsymbol{h} \leq 0$. A to je přesně to, co jsme chtěli.

**Věta XI.2 a její důkaz:**

- Tato věta dává postačující podmínky pro ostré lokální extrémy. Pokud $f$ je třídy $C^2$, $\nabla f(\boldsymbol{a}) = \boldsymbol{o}$ a $\nabla^2 f(\boldsymbol{a})$ je negativně definitní, je v bodě $\boldsymbol{a}$ ostré lokální maximum. Pokud $\nabla f(\boldsymbol{a}) = \boldsymbol{o}$ a $\nabla^2 f(\boldsymbol{a})$ je pozitivně definitní, je v bodě $\boldsymbol{a}$ ostré lokální minimum.

  Případ $n = 1$ byl dokázán v průběhu důkazu Věty XI.1, nicméně obecný případ vyžaduje jiný postup, jak uvidíme níže.

- Součástí věty je i bod (iii), který říká, že pokud $\nabla f(\boldsymbol{a}) = \boldsymbol{o}$ a $\nabla^2 f(\boldsymbol{a})$ je indefinitní, pak v bodě $\boldsymbol{a}$ je sedlový bod.

  To je již snadné, protože to plyne z Věty XI.1. Nechť tedy $\nabla f(\boldsymbol{a}) = \boldsymbol{o}$ a $\nabla^2 f(\boldsymbol{a})$ je indefinitní. Pak $\boldsymbol{a}$ je stacionární bod. Navíc, protože $\nabla^2 f(\boldsymbol{a})$ není pozitivně semidefinitní ani negativně semidefinitní, z Věty XI.1 plyne, že v bodě $\boldsymbol{a}$ není lokální minimum ani lokální maximum.

  Tedy $\boldsymbol{a}$ je stacionární bod, v němž není lokální extrém. To je ovšem přesně definice sedlového bodu.

- Důkaz bodu (i):

  Předpokládejme tedy, že $f$ je třídy $C^2$, $\nabla f(\boldsymbol{a}) = \boldsymbol{o}$ a $\nabla^2 f(\boldsymbol{a})$ je negativně definitní.

  Důkaz provedeme v několika krocích:

**Krok 1:** Označme $\alpha = \max\{\boldsymbol{h}^T\nabla^2 f(\boldsymbol{a})\boldsymbol{h} \colon \boldsymbol{h} \in \mathbf{R}^n, \|h\| = 1\}$. Pak $\alpha < 0$.

Nejprve si uvědomme, že funkce $\boldsymbol{h} \mapsto \boldsymbol{h}^T\nabla^2 f(\boldsymbol{a})\boldsymbol{h}$ (tj. kvadratická forma reprezentovaná maticí $\nabla^2 f(\boldsymbol{a})$) je spojitá na $\mathbf{R}^n$.

Dále, množina

$$S = \{\boldsymbol{h} \in \mathbf{R}^n \colon \|h\| = 1\}$$

je omezená a uzavřená (norma, tj. funkce $h \mapsto \|h\|$, je spojitá funkce), tedy kompaktní (viz Věta V.9).

Protože zřejmě $S \neq \emptyset$, podle věty o nabývání extrémů (Věta V.10) nabývá funkce $\boldsymbol{h} \mapsto \boldsymbol{h}^T\nabla^2 f(\boldsymbol{a})\boldsymbol{h}$ svého minima i maxima. Tedy výše uvedená definice čísla $\alpha$ má smysl – právě jsme ukázali, že maximum existuje.

Navíc platí $\alpha < 0$. Protože $\nabla^2 f(\boldsymbol{a})$ je negativně definitní, je

$$\forall \boldsymbol{h} \in S \colon \boldsymbol{h}^T\nabla^2 f(\boldsymbol{a})\boldsymbol{h} < 0,$$

tedy i maximum je záporné.

**Krok 2:** $\forall \boldsymbol{h} \in S \colon \boldsymbol{h}^T\nabla^2 f(\boldsymbol{a})\boldsymbol{h} \leq \alpha\|\boldsymbol{h}\|^2$.

Pro $\boldsymbol{h} = \boldsymbol{o}$ to zřejmě platí. Pro $\boldsymbol{h} \neq \boldsymbol{o}$ je

$$\boldsymbol{h}^T\nabla^2 f(\boldsymbol{a})\boldsymbol{h} = \|\boldsymbol{h}\|^2 \cdot \underbrace{\frac{\boldsymbol{h}^T}{\|\boldsymbol{h}\|}\nabla^2 f(\boldsymbol{a})\frac{\boldsymbol{h}}{\|h\|}}_{\leq\alpha,\ \text{protože } \left\|\frac{\boldsymbol{h}}{\|h\|}\right\|=1} \leq \alpha\|\boldsymbol{h}\|^2.$$

**Krok 3:** Aplikace Věty X.9 a dokončení důkazu:

Protože $G$ je otevřená množina, existuje $\Delta > 0$, že $B(\boldsymbol{a}, \Delta) \subset G$. Pak $f \in C^2(B(\boldsymbol{a}, \Delta))$, a tedy můžeme použít Větu X.9.

Podle ní existuje funkce $\omega$ definovaná na $B(\boldsymbol{a}, \Delta)$, pro kterou platí $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} \omega(\boldsymbol{x}) = 0$ a

$$f(\boldsymbol{x}) = T_2^{f,\boldsymbol{a}}(\boldsymbol{x}) + \omega(\boldsymbol{x}) \cdot \|\boldsymbol{x} - \boldsymbol{a}\|^2, \quad \boldsymbol{x} \in B(\boldsymbol{a}, \Delta).$$

Použijeme definici Taylorova polynomu a předpoklad, že $\boldsymbol{a}$ je stacionární bod a dostaneme, že pro každé $\boldsymbol{x} \in B(\boldsymbol{a}, \Delta)$ platí

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{a}) + \underbrace{\underbrace{\nabla f(\boldsymbol{a})}_{=\boldsymbol{o}} \cdot (\boldsymbol{x} - \boldsymbol{a})}_{=0} + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T\nabla^2 f(a)(\boldsymbol{x} - \boldsymbol{a}) + \omega(\boldsymbol{x}) \cdot \|\boldsymbol{x} - \boldsymbol{a}\|^2 \\ &= f(\boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T\nabla^2 f(a)(\boldsymbol{x} - \boldsymbol{a}) + \omega(\boldsymbol{x}) \cdot \|\boldsymbol{x} - \boldsymbol{a}\|^2. \end{aligned}$$

Tedy, pro $\boldsymbol{x} \in B(\boldsymbol{a}, \Delta)$ platí

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) = \frac{1}{2}\underbrace{(\boldsymbol{x}-\boldsymbol{a})^T \nabla^2 f(a)(\boldsymbol{x}-\boldsymbol{a})}_{\leq \alpha\|\boldsymbol{x}-\boldsymbol{a}\|^2 \text{ dle Kroku 2}} + \omega(\boldsymbol{x}) \cdot \|\boldsymbol{x}-\boldsymbol{a}\|^2$$

$$\leq \frac{1}{2}\alpha\|\boldsymbol{x}-\boldsymbol{a}\|^2 + \omega(\boldsymbol{x}) \cdot \|\boldsymbol{x}-\boldsymbol{a}\|^2 = \Big(\underbrace{\tfrac{\alpha}{2} + \omega(\boldsymbol{x})}_{\longrightarrow \frac{\alpha}{2}+0=\frac{\alpha}{2}<0}\Big)\|\boldsymbol{x}-\boldsymbol{a}\|^2.$$

Protože $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}}(\frac{\alpha}{2}+\omega(\boldsymbol{x})) = \frac{\alpha}{2} < 0$, podle věty o limitě a nerovnostech (Věta IV.5) existuje $\delta > 0$, že

$$\forall \boldsymbol{x} \in B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\} : \frac{\alpha}{2} + \omega(\boldsymbol{x}) < 0.$$

Proto pro $\boldsymbol{x} \in B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\}$ platí

$$f(\boldsymbol{x}) - f(\boldsymbol{a}) \leq \underbrace{(\tfrac{\alpha}{2} + \omega(\boldsymbol{x}))}_{<0}\underbrace{\|\boldsymbol{x}-\boldsymbol{a}\|^2}_{>0} < 0,$$

tj. $f(\boldsymbol{x}) < f(\boldsymbol{a})$.

To ovšem ukazuje, že v bodě $\boldsymbol{a}$ má funkce $f$ ostré lokální maximum. A důkaz je hotov.

- Bod (ii) se dokáže analogicky, případně aplikací bodu (i) na funkci $-f$.

**Závěrečné shrnutí a příklady:** Obě věty jsou pouze implikace, nikoli ekvivalence. Věta XI.1 dává podmínky nutné a nikoli postačující, Věta XI.2 dává podmínky postačující a nikoli nutné.

Co tyto dvě věty říkají, nyní shrneme:

Nechť $G \subset \mathbf{R}^n$ je otevřená, $f \in C^2(G)$, $\boldsymbol{a} \in G$ je stacionární bod (tj. $\nabla f(\boldsymbol{a}) = 0$).

Pak mohou nastat následující možnosti:

- $\nabla f^2(\boldsymbol{a})$ je negativně definitní.

  Pak je v $\boldsymbol{a}$ ostré lokální maximum (podle Věty XI.2(i)).

- $\nabla f^2(\boldsymbol{a})$ je pozitivně definitní.

  Pak je v $\boldsymbol{a}$ ostré lokální minimum (podle Věty XI.2(ii)).

- $\nabla f^2(\boldsymbol{a})$ je indefinitní (tj. není ani pozitivně semidefinitní ani negativně semidefinitní).

  Pak v $\boldsymbol{a}$ není lokální extrém, je tam sedlový bod (podle Věty XI.2(iii)).

- $\nabla f^2(\boldsymbol{a})$ je negativně semidefinitní a zároveň pozitivně semidefinitní.

  To může nastat jedině v případě, že $\nabla f^2(\boldsymbol{a})$ je nulová matice.

  Pak Věty XI.1 a XI.2 neříkají vůbec nic, a tedy se může stát téměř cokoli, což ilustrujeme následujícími příklady funkcí, z nichž všechny mají v bodě $[0,0]$ nulový gradient i Hessovu matici:

  - Funkce $f(x,y) = 0$ je konstatní, tedy má v $[0,0]$ zároveň lokální minimum i lokální maximum.
  - Funkce $f(x+y) = (x+y)^4$ má v bodě $[0,0]$ lokální minimum, ale ne ostré.
  - Funkce $f(x+y) = -(x+y)^4$ má v bodě $[0,0]$ lokální maximum, ale ne ostré.
  - Funkce $f(x+y) = x^4+y^4$ má v bodě $[0,0]$ ostré lokální minimum.
  - Funkce $f(x+y) = -x^4-y^4$ má v bodě $[0,0]$ ostré lokální maximum.
  - Funkce $f(x+y) = x^4-y^4$ má v bodě $[0,0]$ sedlový bod.

- $\nabla f^2(\boldsymbol{a})$ je negativně semidefinitní, není negativně definitní a není to nulová matice.

  Pak $\nabla f^2(\boldsymbol{a})$ není pozitivně semidefinitní, a tedy v bodě $\boldsymbol{a}$ není lokální minimum (podle Věty XI.1(ii)). Může tam být lokální maximum (ostré nebo neostré) nebo sedlový bod.

  O tom svědčí následující příklady:

  - Funkce $f(x, y) = -x^2 - y^4$ má v bodě $[0, 0]$ ostré lokální minimum.
  - Funkce $f(x, y) = -(x + y)^2$ má v bodě $[0, 0]$ lokální minimum, ale ne ostré.
  - Funkce $f(x, y) = -x^2 + y^4$ má v bodě $[0, 0]$ sedlový bod.

- $\nabla f^2(\boldsymbol{a})$ je pozitivně semidefinitní, není pozitivně definitní a není to nulová matice.

  Pak $\nabla f^2(\boldsymbol{a})$ není negativně semidefinitní, a tedy v bodě $\boldsymbol{a}$ není lokální maximum (podle Věty XI.1(i)). Může tam být lokální minimum (ostré nebo neostré) nebo sedlový bod.

  O tom svědčí následující příklady:

  - Funkce $f(x, y) = x^2 + y^4$ má v bodě $[0, 0]$ ostré lokální maximum.
  - Funkce $f(x, y) = (x + y)^2$ má v bodě $[0, 0]$ lokální maximum, ale ne ostré.
  - Funkce $f(x, y) = x^2 - y^4$ má v bodě $[0, 0]$ sedlový bod.